

本周周报(9.9-9.16):

解聪

本周工作:

1. 淘宝交易地图

周一到周三修对交易图的可视化设计进行了修改。

1.加入了地理信息，采用矢量图数据，以SVG的格式读入程序。采用了一定的透明度绘制，以减少对前景的干扰。

2. 修改了核密度估计的计算方式。

为了解决活动频繁的路径上，所有城市的密度值都较高的问题，对路径中间段的密度值做了衰减。比如北京到上海交易较多，在这两个城市之间路径上密度值累积较大，会导致这样的误解：途径的每个城市的交易都比较密集。

因此，先对路径中间的密度值累积算法进行了修改。路径中的密度值乘上与路径位置相关的系数，使得路径两端的密度值较高，而路径中间的密度值较低。首先尝试系数与位置是线性关系。后来发现改为二次关系效果较好，如图二所示。

这样基本上可以减小路径中，路过的交易对路径的密度值的累积权重。

另外，值得一提的是。通过此种路径方法，我发现了一些有用的信息：

1. 原先被认为是贸易中心的城市，后来发现仅仅是路过比较多。

代表城市：南昌，徐州。图中红圈所示。

这些城市的一般特点是交易路径经过的比较多。调整累积算法后，使得经过的交易对城市累积减小。因此这些城市的密度值相应减小

2. 原先被认为交易量不大的城市，在调整之后在整个密度场中权值加大。反映出是该局部地区的贸易中心。

代表城市：昆明，太原。图中紫圈所示

这些城市的一般特点是总量在全局看来虽然不大，但是在局部是一个交易量的极值。调整累积算法后，使得该处的权值增加。

3. 一些二级城市（未标注的算二级）的交易量高过许多一级城市（有标注的算一级）。

代表城市：厦门，温州。图中绿圈所示。

由于算法的调整使得大部分路径中的密度值减小，因此一些交易量大的二级城市（未标注）也被凸显出来。厦门虽然在全局看来交易量不大，但是局部看交易量还是很大。

但是因为路径数据是模拟所得，所以以上结论是否正确还是有待考证。

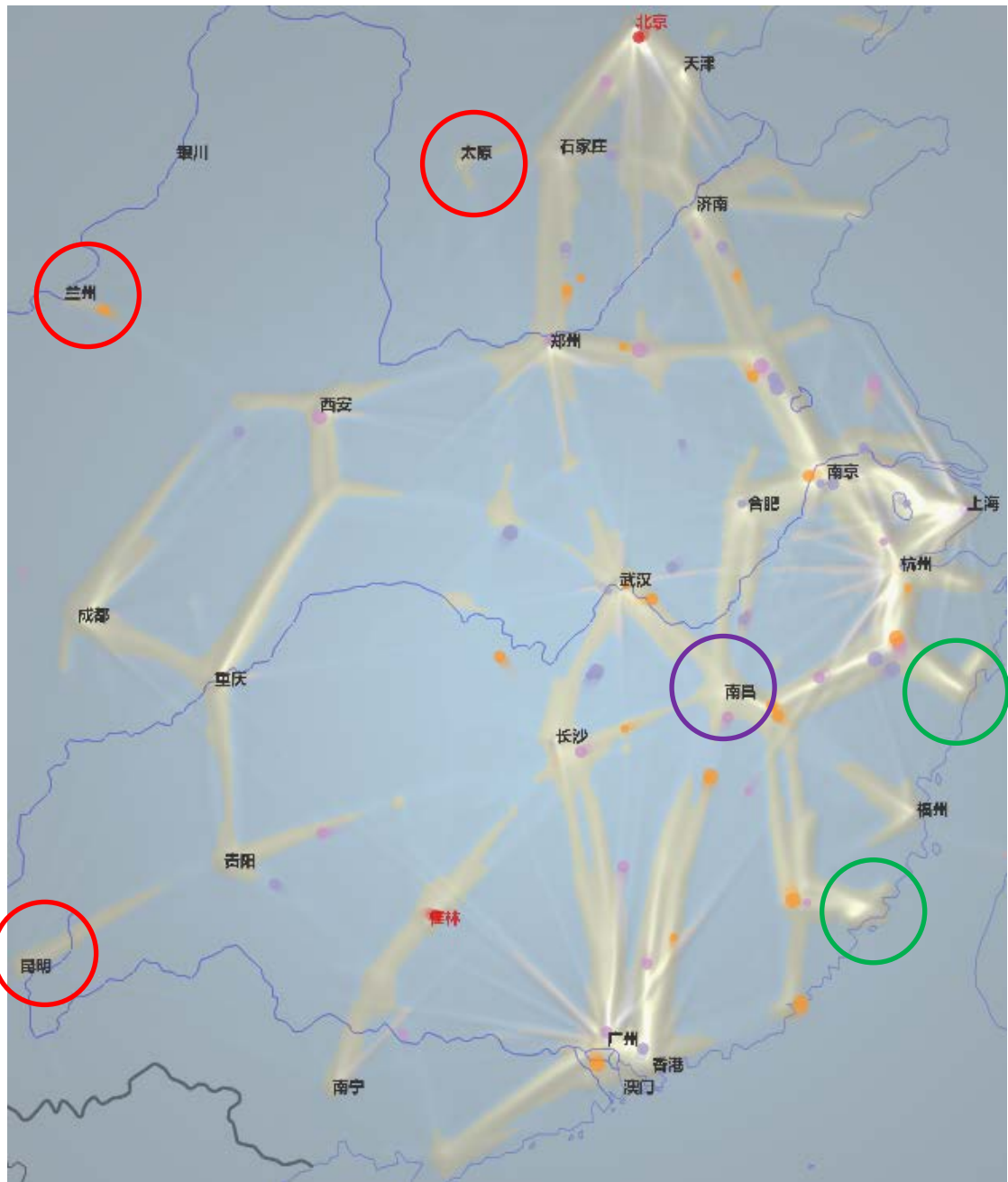


图1.采用了二次的路径位置与密度场衰减的关系。本图中可以明显看京沪线的中段密度场有所减弱。另外采用了这样的效果就可以看出一些城市的辐射形的交易，比如广州，杭州。圈中城市的信息的变化与图2对比可得。

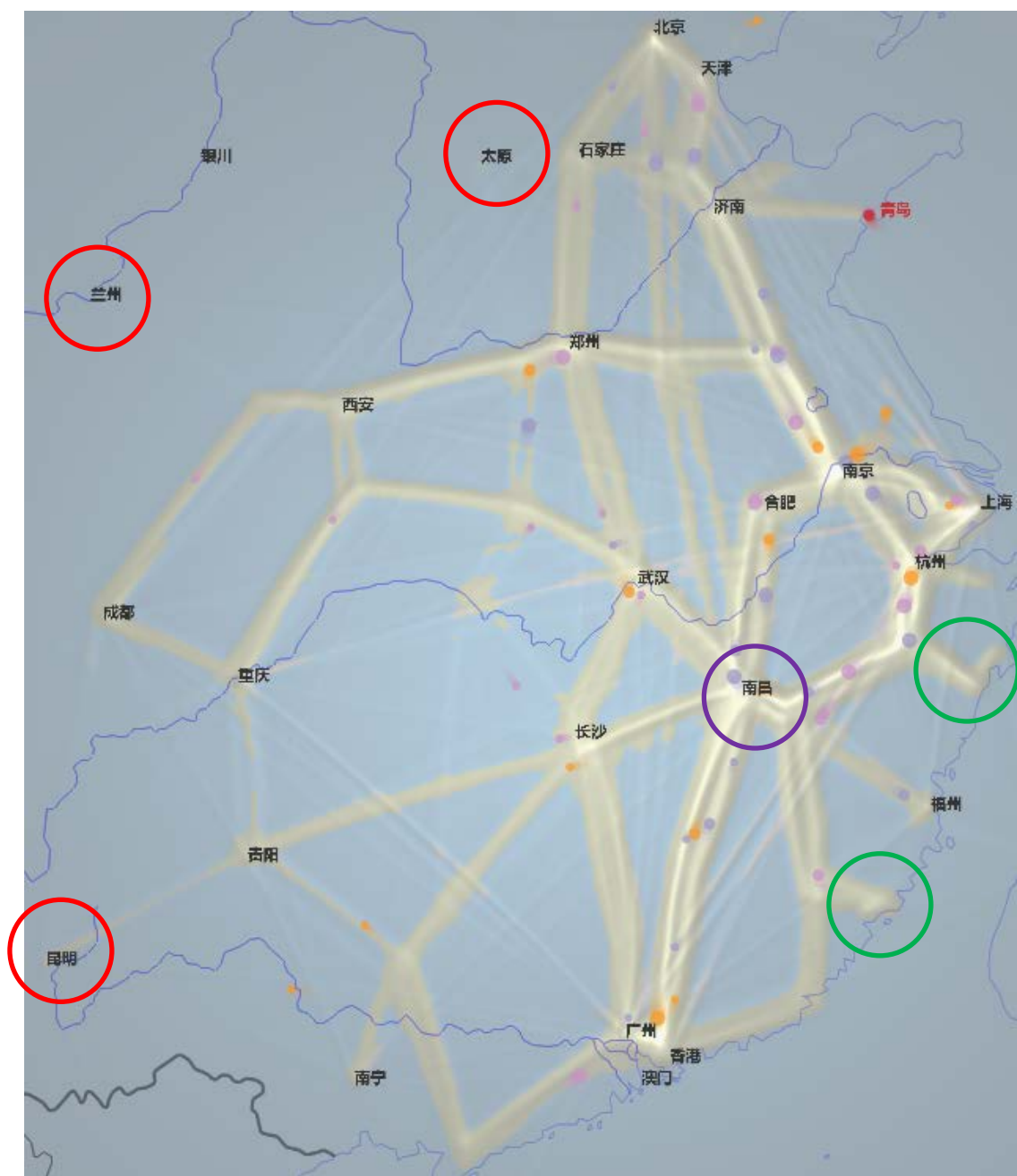


图2.先前的无衰减效果，与图一作对比。可以发现红圈与绿圈城市的交易在此图中较小，但是在图一中较大，说明该地不是交通枢纽，而是局部贸易中心。紫色圈城市在该图中较大，在上图中较小，可以发现这些城市是交通枢纽。

2. DataV 组件

开发了饼状图，目前还在持续调试阶段。

3. 准备下周的组会论文报告

下周工作：

1. 淘宝交易地图

对下一步的工作，以及程序的设计，做一个计划。

2. DataV 组件的开发

完成饼图的组件。

3. 教材